第 37 卷第 6 期
录用定稿
计算机应用研究
Application Research of Computers
ChinaXiv合作期刊
Vol. 37 No. 6
Accepted Paper

基于 PN 码随机化的 MSAC 攻击防御方法 *

许新忠 [1]，张连成 [2†]，燕菊维 [2]

(1. 河南艺术职业学院，郑州 450011; 2. 数学工程与先进计算国家重点实验室，郑州 450001)

摘　要：基于流速率的直序扩频（DSSS）流水印技术在嵌入多比特的水印信息时采用同一伪随机（PN）码，使得已标记数据流具有自相似性，均方自相关（MSAC）攻击通过单条已标记数据流就可检测 DSSS 流水印的存在性，严重破坏了 DSSS 流水印的隐蔽性。PN 码正交化方法虽可消除已标记数据流的自相关性，进而抵御 MSAC 攻击的检测，但正交 PN 码难以生成，应用范围受限。为此，提出基于 PN 码随机化的 MSAC 攻击防御方法，在向目标数据流嵌入每个水印位时均采用随机选择的不同长度的 PN 码进行扩展，使得已嵌入 DSSS 流水印的数据流速率的均方自相关不再呈现周期性峰值，进而可在 MSAC 攻击面前保持隐蔽性。理论分析与实验结果表明，基于 PN 码随机化的 MSAC 攻击防御方法可有效抵御 MSAC 攻击的检测，且所采用的 PN 码易于生成，便于部署和应用。

关键词：流水印；直序扩频；均方自相关攻击；PN 码正交化；PN 码随机化

中图分类号：TP309.02　　doi: 10.19734/j.issn.1001-3695.2018.11.0891

Pn code randomization based msac attack resistance method

Xu Xinzhong[1], Zhang Liancheng[2], Yan Juwei[2]

(1. *Henan Vocational College of Art*, *Zhengzhou* 450011, *China*; 2. *State Key Laboratory for Mathematical Engineering & Advanced Computing*, *Zhengzhou* 450001, *China*)

Abstract: The flow rate based direct sequence spread spectrum (DSSS) flow watermarking technique used the same pseudo-random (PN) code to modulate the multi-bit watermark signal, as a result, the DSSS watermarked flow had self-similarity, and the mean-square autocorrelation (MSAC) attack could detect the existence of the DSSS flow watermark by using single DSSS watermarked flow, which seriously destroyed the stealthiness of the DSSS flow watermark. Although the PN code orthogonalization method could eliminate the autocorrelation of the DSSS watermarked flow and thus resisted the MSAC attack, however, the orthogonal PN codes were difficult to generate, which limit its application range. Therefore, this paper proposed the PN code randomization based MSAC attack resistance (PNCR-MAR) method, when the watermark signal was embedded into the target flow, the PN codes of different lengths, randomly selected from a PN code set, were used to spread each watermark bit, so that the mean-square autocorrelation of the flow rate of the watermarked flow no longer exhibited periodic peaks, as a result, this PNCR-MAR method could defend against the MSAC attack. Theoretical analysis and experimental results show that this proposed PNCR-MAR method can effectively resist the MSAC attack, and the adopted PN codes are easy to generate, which is suitable for deployment and application.

Key words: flow watermarking; direct sequence spread spectrum; mean-square autocorrelation attack; pseudo-noise code orthogonalization; pseudo-noise code randomization

0　引言

与被动流相关技术相比，主动流水印技术通过主动调制发送者数据流特征（如对特定数据包进行延迟或影响数据流速率等）进而嵌入水印信息来帮助确认发送者和接收者的通信关系[1~3]，在跳板攻击追踪和匿名用户关联等方面具有准确率高、误报率低等优点。典型的流水印技术有基于时间间隔的流水印（interval-based watermarking，IBW）[4,5]、基于间隔重心的流水印（interval centroid based watermarking，ICBW）[6]、基于间隔到达时延（inter-packet delay，IPD）的流水印[7]、RAINBOW 流水印[8,9]和 SWIRL 流水印[10]等，分别通过调制时间间隔、间隔重心、间隔到达时延等流量特征来嵌入水印信息。然而，这些流水印技术都需要对目标数据流拥有较为完全的控制权才能进行水印信息的嵌入[11]。

相比而言，基于直序扩频（direct sequence spread spectrum，DSSS）的流水印技术[12]不需对目标数据流具备完全控制权，通过调制目标数据流的速率即可嵌入相应水印信息，可有效追踪恶意匿名用户和跳板攻击源。

然而，DSSS 流水印技术在向目标数据流嵌入水印信息时，对于不同的水印位采用同一 PN（pseudo-noise，伪噪声）码进行嵌入，导致嵌入 DSSS 流水印信息的已标记数据流存在自相似性，其速率的时间序列的均方自相关（mean-square autocorrelation，MSAC）呈现周期性峰值，Jia 等人[13,14]据此提出 MSAC 攻击方法，根据单条已嵌入 DSSS 流水印的已标记数据流即可检测出其中 DSSS 流水印的存在性，即使对 DSSS 流水印所用的 PN 码毫无所知，且不需数据流同步[15]。

为提升 DSSS 流水印技术的 MSAC 攻击防御能力，对比分析现有 MSAC 攻击防御方法的缺点和不足，研究 MSAC

收稿日期：2018-11-26；**修回日期：**2019-01-28　**基金项目：**国家自然科学基金资助项目（61402526，61402525）

作者简介：许新忠（1978-），男，河南修武人，讲师，硕士，主要研究方向为信息安全；张连成（1982-），男（通信作者），副教授，博士，主要研究方向为流量追踪、IPv6 网络安全、软件定义网络安全（liancheng17@gmail.com）；燕菊维（1984-），女，讲师，硕士，主要研究方向为网络态势感知（yan_jvwei@163.com）.

攻击防御新方法，提出基于 PN 码随机化的 MSAC 攻击防御（PN code randomization based MSAC attack resistance，PNCR-MAR）方法，并对 PNCR-MAR 方法的抗 MSAC 攻击能力进行理论分析和实验验证。

1 MSAC 攻击防御研究现状分析

1.1 MSAC 攻击

Jia 等人[13,14]提出 MSAC 攻击方法来检测 DSSS 流水印的存在性。由于 DSSS 流水印技术在调制多比特的水印信息时采用同一个 PN 码，使得已标记数据流具有自相似性，使得攻击者通过单条已标记数据流的流量速率时间序列的均方自相关即可检测 DSSS 流水印的存在性，即使对其所用的 PN 码毫无所知。

该攻击方法复杂度低，比基于多流攻击（multi-flow attack，MFA）的 DSSS 流水印检测方案[16,17]更加灵活和准确，对 DSSS 流水印的隐蔽性造成巨大挑战。

1.2 现有 MSAC 攻击防御方法分析

为抵御 MSAC 攻击，张连成等人提出可抵御 MSAC 和多流攻击的扩频流水印（MSAC and multi-flow attacks resistant spread spectrum watermarking，MMAR-SSW）方案[18]和基于时间间隔的扩频流水印（interval-based spread spectrum watermarking，IBSSW）技术[19,20]，使用多个正交 PN 码来扩展水印信息的不同水印位，并使用相同的正交 PN 码来解扩以恢复水印信息。由于这些 PN 码是正交的，因此已嵌入水印的已标记数据流速率的时间序列的均方自相关就不会显现出周期性峰值，进而可有效抵御 MSAC 攻击。

在抵御 MSAC 攻击的检测时，MMAR-SSW 方案和 IBSSW 技术本质上采用的是 PN 码正交化方法。该方法的 MSAC 攻击防御效果较好，然而，要求在对水印位进行扩展时采用正交 PN 码，但是正交 PN 码并不容易获得，限制了该方法的应用。

张璐等人[21]提出基于时隙质心流水印的匿名通信追踪技术，由于时隙组是随机分配的，攻击者在不确定时隙组具体分配方式的情况下难以得到相应码片所具体对应的时隙质心，进而无法判断流是否在短时间内出现了自相似性，因此对 MSAC 攻击具有免疫力，然而，文献[21]中只进行了文字论述，未进行理论推导与实验验证。不过，其所采取的随机化思想值得借鉴和学习。

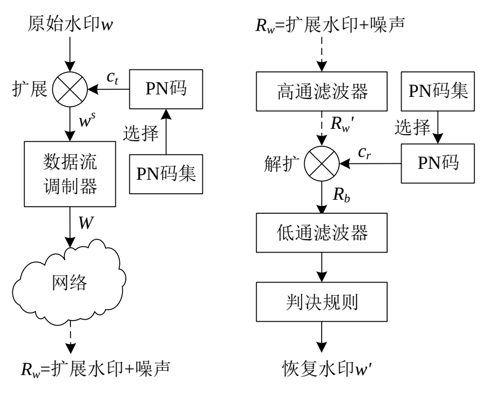
1.3 基于 PN 码随机化的 MSAC 攻击防御新思路

针对 PN 码正交化方法的不足，本文提出“一比特一 PN 码”的 PN 码随机化思想，在向目标数据流嵌入每个水印位时均采用随机选择的、长度不同的 PN 码进行扩展，使得水印信息经由 PN 码扩展后呈现出随机性，已嵌入流水印的已标记数据流速率的均方自相关不再呈现周期性峰值，进而可在 MSAC 攻击面前保持隐蔽性，而且此方法下对 PN 码没有特殊要求，其生成和使用都很容易，应用更为广泛。

2 基于 PN 码随机化的 MSAC 攻击防御方法

基于 PN 码随机化的 MSAC 攻击防御方法的基本框架如图 1 所示，由水印嵌入器和水印检测器两个部分组成。水印嵌入器负责对原始水印进行扩展，并调制目标数据流，已标记数据流最终被送入网络中进行传输，可能经过匿名通信节点、跳板链等的处理和传递。水印检测器的功能是从被干扰后的已标记数据流中检测、恢复出其中嵌入的水印信息，如果恢复出的水印信息与原始水印相同，那么就可确认发送者和接收者之间的通信关系。

下面分别对水印嵌入器的水印信息扩展和嵌入过程、水印检测器的水印信息检测过程与判决规则进行介绍。假设下面的参数是水印嵌入器和水印检测器共享的（通过某种秘密途径协商或预先设定等）：码片时长 T_c、l 位水印 $w=\{w_0,w_1,\cdots,w_{l-1}\}$、PN 码集（包含多个长度不同的 PN 码）$PN$、PN 码选取随机值 s。



(a)水印嵌入器 (b)水印检测器

(a) Watermark embedder (b) watermark detector

图 1 PNCR-MAR 方法处理流程

Fig. 1 PNCR-MAR method and processing procedure

2.1 水印信息扩展和嵌入过程

水印嵌入器采取以下步骤进行水印信息的扩展和嵌入：

a)依据 PN 码选取随机值 s 从 PN 码集中为每个水印位 w_i（‘+1’或者‘-1’，$0\le i\le l-1$）随机选取 PN 码 $c_{t,i}$（长度为 $N_{c,i}$）。

b)依次依据选取的 PN 码 $c_{t,i}$ 对水印位 w_i 进行扩展（即对不同的水印位采用不同的 PN 码，长度为 $N_{c,i}$，如图 2 所示），可得待传输的扩展水印 w_i^s，w_i^s 可被表示为

$$w_i^s = w_i c_{t,i} \tag{1}$$

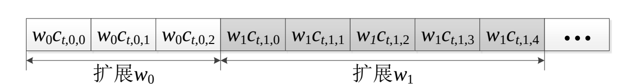


图 2 PN 码随机化的水印位扩展

Fig. 2 Watermark bit spreading of the PN code randomization method

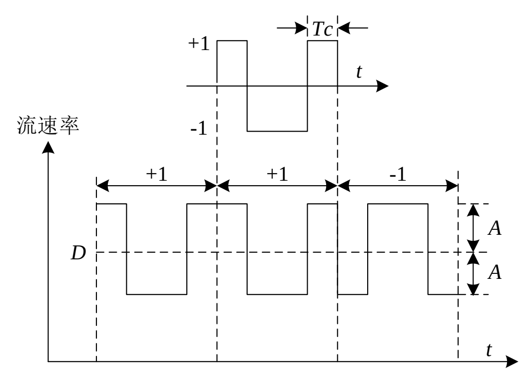


图 3 PNCR-MAR 方法数据流调制

Fig. 3 Flow modulation of the PNCR-MAR method

c)将 w_i^s 嵌入目标数据流 f^u 中。当扩展水印位为‘+1’时，对目标数据流施加弱干涉，因此数据流在 T_c 秒时间段内的速率较大。当扩展水印位为‘-1’时，对目标数据流施加强干涉，因此数据流在 T_c 秒时间段内的速率较小。假设数据流的平均速率为 D，高速间隔的速率为 $D+A$，低速间隔的速率为 D-A（其中 A 为调制幅度），如图 3 所示。因此嵌入水印后的数据流速率 W 为

$$W = Awc_t + D \tag{2}$$

d)重复上述步骤 b)c)来嵌入 l 位的水印信息。

e)将已标记数据流发送到网络上进行传输。

2.2 水印信息检测过程

已标记数据流经过网络传输、匿名处理等干扰之后，流经放置在接收端附近的水印检测器。水印检测器采取以下步骤进行水印信息的检测：

a)水印检测器捕获到已标记数据流 f^w 之后，将其分为长度为 T_c 的时间段，计算每个段的平均速率，依据 PN 码选取随机值 s 从 PN 码集中为每个水印位选取对应的 PN 码（获取 PN 码本身及对应的 PN 码长度），然后计算与所选取 PN 码长度对应的连续时间段的平均速率。假设网络中的噪声为 ξ，那么水印检测器接收到的受干扰后的已标记数据流的速率信号 R_w（对于单个水印位 w_i 而言）为

$$R_w = Aw_i c_{t,i} + D + \xi \tag{3}$$

b)将接收到的 R_w 经过一个高通滤波器滤去分量 D，那么过滤后的信号 R_w' 为

$$R_w' = Aw_i c_{t,i} + \xi \tag{4}$$

c)获取 R_w' 后，依据所选取的对应 PN 码 $c_{r,i}$ 计算信号 R_b：

$$R_b = Aw_i c_{t,i} \cdot c_{r,i} + \xi \cdot c_{r,i} \tag{5}$$

d)再使用低通滤波器滤除高频噪声，根据 2.3 节给出的判决规则就可获得恢复水印信息位 w_i'。

e)重复步骤 a)~d)即可恢复得到 l 位的水印信息 w'。

f)如果恢复水印 w' 与原始水印 w 相同，则水印检测器报告发现水印信息，那么发送者和接收者的关系即可确定。

可以看到，在 PNCR-MAR 方法中，原始水印长度 l、码片时长 T_c、PN 码长度 $N_{c,i}$ 和水印调制幅度 A 都会影响数据流追踪效果。

2.3 判决规则

为便于分析，暂时不考虑噪声等干扰，水印检测器接收到的已标记数据流的速率信号 R_w（对于单个水印位 w_i 而言）为

$$R_w = Aw_i c_{t,i} + D \tag{6}$$

假设 $t = Aw_i c_{t,i}$，对信号 R_w 进行离散傅里叶变换可得

$$R(k) = \sum_{n=0}^{N_{c,i}-1} (t(n)+D) W_{N_{c,i}}^{kn} = T(k) + \sum_{n=0}^{N_{c,i}-1} D W_{N_{c,i}}^{kn} \tag{7}$$

$$r(n) = \frac{1}{N_{c,i}} \sum_{k=0}^{N_{c,i}-1} R(k) W_{N_{c,i}}^{-kn} \tag{8}$$

其中: $W_{N_{c,i}} = e^{-j(2\pi/N_{c,i})}$，由式(7)可得

$$\begin{aligned} R(0) &= \sum_{n=0}^{N_{c,i}-1} (t(n)+D) W_{N_{c,i}}^{0n} = \sum_{n=0}^{N_{c,i}-1} Aw_i c_{t,i}(n) + N_{c,i} D \\ &= Aw_i \sum_{n=0}^{N_{c,i}-1} c_{t,i}(n) + N_{c,i} D \end{aligned} \tag{9}$$

假设 $\sum_{n=0}^{N_{c,i}-1} c_{t,i}(n) = x$，则有

$$R(0) = Aw_i x + N_{c,i} D \tag{10}$$

利用高通滤波器过滤掉 $R(0)$ 后，可得

$$\begin{aligned} r'(n) &= \frac{1}{N_{c,i}} \left[\sum_{k=0}^{N_{c,i}-1} R(k) W_{N_{c,i}}^{-kn} - R(0) \right] \\ &= \frac{1}{N_{c,i}} \left[\sum_{k=0}^{N_{c,i}-1} \left(T(k) + \sum_{m=0}^{N_{c,i}-1} D W_{N_{c,i}}^{-km} \right) W_{N_{c,i}}^{-kn} - R(0) \right] \\ &= \frac{1}{N_{c,i}} \left[\sum_{k=0}^{N_{c,i}-1} T(k) W_{N_{c,i}}^{-kn} + N_{c,i} D - Aw_i x - N_{c,i} D \right] \\ &= \frac{1}{N_{c,i}} \left[\sum_{k=0}^{N_{c,i}-1} T(k) W_{N_{c,i}}^{-kn} - Aw_i x \right] \\ &= t(n) - Aw_i x / N_{c,i} \end{aligned} \tag{11}$$

因此可知

$$\begin{aligned} \sum R_b / N_{c,i} &= \frac{\sum (r' c_{r,i})}{N_{c,i}} = \frac{\sum [(t - Aw_i x / N_{c,i}) c_{r,i}]}{N_{c,i}} \\ &= \frac{\sum [(Aw_i c_{t,i} - Aw_i x / N_{c,i}) c_{r,i}]}{N_{c,i}} \\ &= \frac{Aw_i N_{c,i} - Aw_i x / N_{c,i} \sum c_{r,i}}{N_{c,i}} \\ &= Aw_i - \frac{Aw_i x^2}{N_{c,i}^2} = Aw_i \left(1 - \frac{x^2}{N_{c,i}^2} \right) \end{aligned} \tag{12}$$

一般情况下，PN 码都是由 m 序列发生器产生，因此，$|x| < |N_{c,i}|$，$1 - x^2 / N_{c,i}^2 > 0$。当 $w_i = +1$ 时，$\sum R_b / N_{c,i} > 0$，当 $w_i = -1$ 时，$\sum R_b / N_{c,i} < 0$，因此，判决规则如下：

$$w_i' = \begin{cases} +1, & \sum R_b / N_{c,i} \ge 0 \\ -1, & \sum R_b / N_{c,i} < 0 \end{cases} \tag{13}$$

3 MSAC 攻击防御能力分析

由随机化的 PN 码所调制的已标记数据流在频域和时域中都显示出类似白噪声的模式，MSAC 攻击者难以在频域和时域中检测到流水印的存在性。

MSAC 攻击的有效性所基于的事实是，DSSS 流水印重复使用相同的 PN 码来扩展信号（即水印）的每个水印位。而在 PNCR-MAR 方法中，每个信号位由不同的PN码来扩展，这些 PN 码不但码值不一样，且长度也不同。

本章分析中涉及到的各符号的含义与第 2 章描述的相同，假设信号位 w_i 和 w_j $(i \ne j)$ 是独立的。调制后的水印信号可以写成

$$\begin{aligned} w^s &= (w_0 c_{t,0}, w_1 c_{t,1}, \cdots, w_{l-1} c_{t,l-1}) \\ &= (w_0 c_{t,0,0}, \ldots w_0 c_{t,0,N_{c,0}-1}, w_1 c_{t,1,0}, \ldots, w_1 c_{t,1,N_{c,1}-1}, \\ & \quad w_{l-1} c_{t,l-1,0}, \ldots, x_{l-1} c_{t,l-1,N_{c,l-1}-1}) \end{aligned} \tag{14}$$

由于 $w_i c_{t,i,j}$ 是独立且相同分布的，因此 $P(w_i c_{t,i,j} = A) = 1/2$ 和 $P(w_i c_{t,i,j} = -A) = 1/2$，因此 $E(w_i c_{t,i,j}) = 0$ 和标准偏差 $\sigma = A$。以下公式可用于估计由 w^s 表示的时间序列的自相关：

$$r(\gamma) = 1/(l-\gamma) \sum_{i=0}^{l-1-\gamma} (a_{i_j} a_{i_{j+\gamma}}) \tag{15}$$

其中：γ 是时移，$a_{i_j} = w_i c_{t,i,j}$ 是 w^s 的第 i 项，且 $i \in [0, l-1]$，$j \in [0, N_{c,i} - 1]$。

MSAC 攻击方法通过计算 $E(r^2(\gamma))$ 来检测 DSSS 流水印的存在性。$r^2(\gamma)$ 是扩展信号 w^s 和具有时移 γ 的信号 w^s 的平方自相关。通过计算 $E(r^2(\gamma))$，将出现周期为 l 的周期性峰值。

当 $\gamma = 0$ 时，从式(14) (15)可知

$$r(0) = \frac{1}{l} \sum_{i=0}^{l-1} (w_i c_{t,i,j} w_i c_{t,i,j}) = \frac{1}{j} \sum w_i^2 c_{t,i,j}^2 \tag{16}$$

因 $w_i^2 = A^2$，$c_{t,i}^2 = 1$，那么有 $r(0) = A^2$ 和 $E(r(0)) = A^4$。

当 $\gamma \ne 0$ 时，从式(14)(15)可知

$$\begin{aligned} r(\gamma) &= \frac{1}{l-\gamma} \sum_{i=0}^{l-1-\gamma} (a_{i_j} a_{i_{j+\gamma}}) \\ &= \frac{1}{l-\gamma} (w_0 c_{t,0} w_\gamma c_{t,\gamma} + \ldots + w_{l-1-\gamma} c_{t,l-1-\gamma} w_{l-1} c_{t,l-1}) \end{aligned} \tag{17}$$

$c_{t,i}$ 是 PN 码集所选取的第 i 个 PN 码，用于扩展第 i 个水印位，因此有 $c_{t,i} \ne c_{t,j}$（如果 $i \ne j$）。当 $\gamma \ne 0$ 时，$E(c_{t,i} c_{t,i+\gamma}) \approx 0$，因此，有 $E(w_i c_{t,i,j} w_{i+\gamma} c_{t,i+\gamma,j}) = E(w_i w_{i+\gamma}) E(c_{t,i,j} c_{t,i+\gamma,j}) = 0$ 和 $r(\gamma) = 0$

（$\gamma \neq 0$）。因此，$E(r^2(\gamma))=0$（$\gamma \neq 0$）。

由上可知，$E(r^2(\gamma))$ 的均值为

$$E(r^2(\gamma)) \approx \begin{cases} A^4, \gamma = 0 \\ 0, \gamma \neq 0 \end{cases} \tag{18}$$

式(18)表明本文所提的 PNCR-MAR 方法所调制的数据流中不会出现周期性峰值，只在 $\gamma=0$ 时显示出一个峰值，这与在文献[12]中所提出的基于同一 PN 码的 DSSS 流水印技术不同，对于后者，MSAC 攻击方法可发现和检测到目标数据流中所嵌入的 DSSS 流水印的自相似性，然而，由于 PNCR-MAR 方法所调制的目标数据流中不会出现周期性峰，因此，PNCR-MAR 方法在追踪业务流时的隐蔽性较好。

4　MSAC 攻击防御实验结果与分析

为测试 PNCR-MAR 方法的 MSAC 攻击防御能力，搭建 MSAC 攻击防御实验环境。

4.1　MSAC 攻击防御实验环境搭建

由于 DSSS 流水印技术采用 NS-2 网络模拟器进行实验测试，为保持可比性，本文也采用 NS-2 网络模拟器来搭建测试环境来测试 PNCR-MAR 方法的 MSAC 攻击防御能力。实验中使用在互联网占主导地位的 TCP 数据流[22]进行测试。

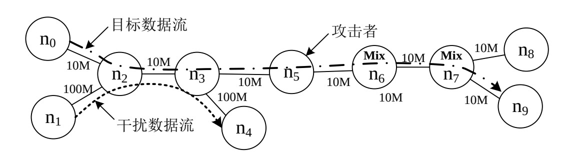


图 4　NS-2 模拟测试环境

Fig. 4　NS-2 simulation test environment

图 4 给出了 NS-2 模拟器中所采用的拓扑结构。其中节点 n_6 和节点 n_7 为 Mix 节点，其所使用的批处理策略有简单代理（simple proxy）、计时 mix（timed mix）和 stop-and-go 或连续 mix，如表 1 所示。

表 1　mix 节点的批处理策略

Table 1　Batch processing strategies for mix nodes

序号	名称	可调参数	算法
S_0	简单代理	没有	没有批处理（或包乱序）
S_1	计时 mix	$<t>$	如果计时器周期 t 过期，就将队列中所有数据流发出
S_2	stop-and-go mix	$<\mu,\sigma^2>$	为每个数据流分配一个延迟（期限），满足均值 μ 和方差 σ^2 的分布，当期限到达就将该数据包发出

在实验时，目标 FTP 数据流从节点 n_0 流到节点 n_9。干扰器使用恒定比特率（constant bit rate，CBR）的 UDP 数据流来调制目标 FTP 数据流以嵌入水印信息，其中 CBR 数据流从节点 n_1 到节点 n_4，与目标 FTP 数据流在节点 n_2 和节点 n_3 之间共享同一链路。根据 TCP 流控制机制，当 CBR 流速率增加时，FTP 流速率就会降低，而当 CBR 流速率降低（如没有 CBR 流量的干扰）时，FTP 流速率就会增加。节点 n_5 为攻击者节点，该节点负责收集已标记数据流，并进行 MSAC 攻击。

测试中，当被 PN 码扩展过后的扩展水印信息的一个码片是‘+1’时，节点 n_1 到节点 n_4 的 CBR 流量就关闭，当该码片是‘-1’时，节点 n_1 到节点 n_4 的 CBR 流量就开启。开启和关闭的时间长度与码片时长保持一致。这样就可使用 CBR 流量来调制目标 FTP 数据流的速率来嵌入水印信息。

为从已标记数据流中恢复出其中嵌入的原始水印信息，必须首先获得 FTP 流速率的时间序列。根据奈奎斯特—香农采样理论，为恢复扩展水印信息，采样周期应比码片时长的一半还要小。

实际中，由于互联网的复杂性和动态性，发送者和接收者之间的时延很难准确预测。实验中，为同步发送者和接收者之间的 PN 码，使用延迟的近似估计值，与数据流路径上的累计传播和排队延迟相近。接着使用基于匹配滤波器的方法搜寻特定范围内的最佳匹配，实验中所使用的范围为[-0.5, 0.5] s。

4.2　检测率测试

因为 PNCR-MAR 方法主要是从 PN 码随机化(不但数值随机，长度也不同)的角度对原始 DSSS 流水印技术做优化，本节主要测试不同的 PN 码平均长度对检测率（检测率是指从目标数据流中正确检测出其中所嵌入的水印信息的比率）的影响。由图 5 可看到，检测率随着 PN 码平均长度的增长而增加，较长的 PN 码平均长度可以获得较高的检测率，因此，可以使用较长的 PN 码来对抗嘈杂网络环境对流水印的影响而获得更好的追踪效果。

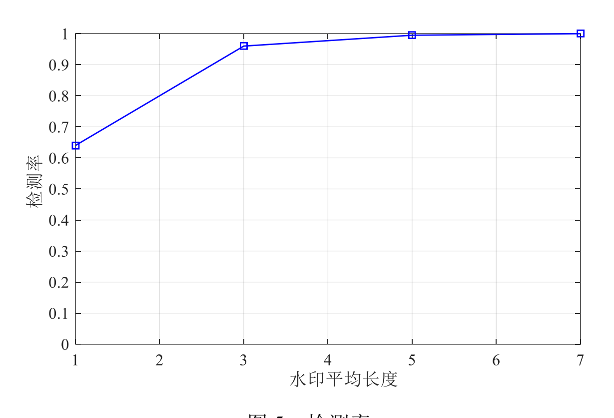


图 5　检测率

Fig. 5　Detection rate

4.3　误报率测试

为测试 PNCR-MAR 方法在应用于 DSSS 流水印技术之后对于 DSSS 流水印技术的误报率（误报率是指从目标数据流中错误检测出其中所嵌入的水印信息的比率）造成的影响，本实验使用码片时长为 0.5 s，PN 码平均长度为 7。DSSS 流水印技术误报率的理论值为 2^{-l}，其中 l 为水印长度[11]。

PNCR-MAR 方法在不同水印长度情况下的误报率对比如图 6 所示，误报率随着水印长度呈指数级降低，符合文献[12]中的分析结果，且 PNCR-MAR 方法与原始 DSSS 流水印技术的误报率相近，因此，PNCR-MAR 方法在抵御 MSAC 攻击的同时，并没有增加 DSSS 流水印技术的误报率。

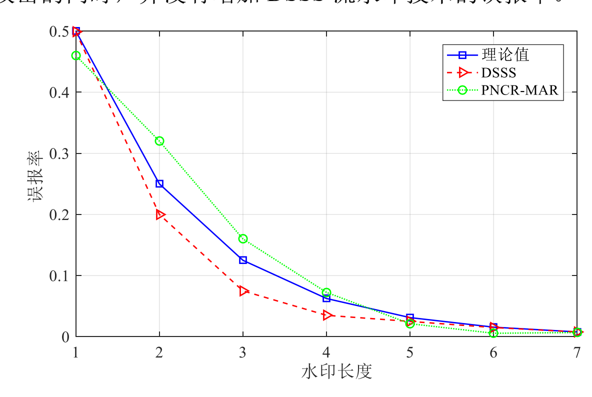


图 6　误报率对比

Fig. 6　Comparison of false positive rates

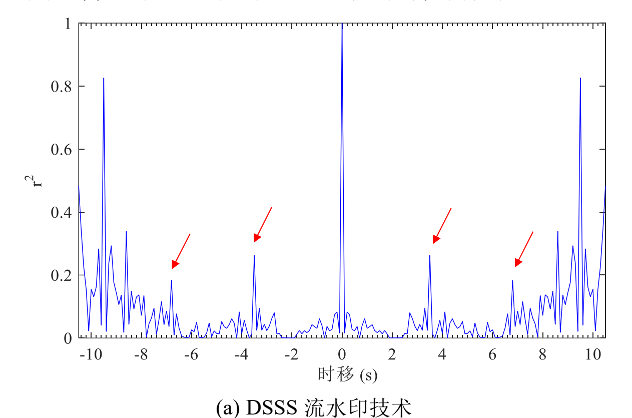
4.4　MSAC 攻击防御能力测试

本节测试 PNCR-MAR 方法的 MSAC 攻击防御能力。实

验中 PN 码平均长度为 7，水印长度为 7，干扰数据流的 CBR 数据流速率为 1.2 Mbps，码片时长为 0.5 s，采样间隔为 0.1 秒。为对比 DSSS 流水印技术和 PNCR-MAR 方法在 MSAC 攻击面临的隐蔽性，计算嵌入水印信息后的已标记数据流在不同时移情况下的均方自相关并进行归一化（r^2）处理后，如图 7 所示。

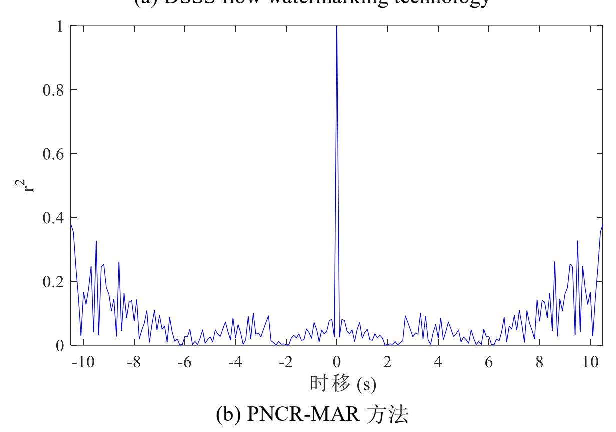
如图 7(a)所示，采用 DSSS 流水印技术嵌入水印信息的已标记数据流的流速率时间序列的均方自相关确实出现了周期性峰值，故 MSAC 攻击可有效检测到 DSSS 流水印的存在性。

如图 7(b)所示，使用 PNCR-MAR 方法嵌入水印信息的已标记数据流的流速率时间序列的均方自相关没有出现周期性峰值，这是因为 PNCR-MAR 方法采用 PN 码随机化思路，对于不同的水印位采用不同的 PN 码进行扩展，消除了已标记数据流的自相关性，故而其流速率的均方自相关并不出现周期性峰值，因此可抵御 MSAC 攻击的检测发现。



(a) DSSS 流水印技术

(a) DSSS flow watermarking technology



(b) PNCR-MAR 方法

(b) PNCR-MAR method

图 7　归一化均方自相关估计

Fig. 7　Estimation of normalized mean-square autocorrelation

此外，为测试 PNCR-MAR 方法对 MSAC 攻击的隐蔽性，使用检测率和误报率作为衡量标准，并使用受试者工作特征（receiver operating characteristic，ROC）曲线（以检测率为横坐标，以误报率为纵坐标绘制的曲线）表示二者之间的关系。当攻击者检测目标数据流是否含有 DSSS 流水印时，他希望是低误报率的同时，获得高检测率。理想情况下，为抵御 MSAC 攻击的检测，应使得检测率和误报率一样高。

如图 8 所示，可以看到 PNCR-MAR 方法下的 ROC 曲线与理想的 ROC 曲线较为吻合，随着检测率的增加，误报率趋向于线性增长，体现了 PNCR-MAR 方法的 MSAC 攻击防御能力，这使得 MSAC 攻击者的检测结果是没有价值的。

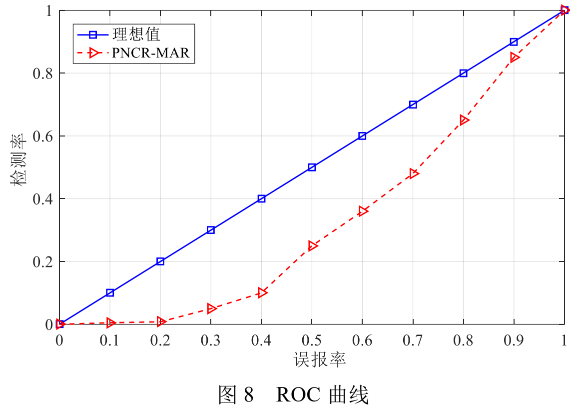


图 8　ROC 曲线

Fig. 8　ROC curves

5　结束语

针对 PN 码正交化方法存在的正交 PN 码难以获取的问题，为提高 DSSS 流水印技术的 MSAC 攻击防御能力，提出基于 PN 码随机化的 MSAC 攻击防御方法。理论分析及实验结果表明，PNCR-MAR 方法可有效抵御 MSAC 攻击的检测发现，且克服了 PN 码正交化的应用难题。

后继拟研究 PN 码随机化方法在抵御多流攻击方面的有效性。此外，Luo 等人[23]提出基于序贯概率比检验（sequential probability ratio testing，SPRT）的 DSSS 流水印检测方法，还提出基于 TCP 流控制机制的 DSSS 流水印移除方法，使得终端用户可有效移除数据流中的 DSSS 流水印，且不需要中间路由器、代理或中继节点的支持。对于该检测和移除方法，目前尚无有效防御手段，后继拟研究相应的防御手段，进一步提高 DSSS 流水印技术的隐蔽性。

参考文献：

[1] Iacovazzi A, Elovici Y. Network flow watermarking: a survey [J]. IEEE Communications Surveys & Tutorials. 2017, 19(1): 512-530.

[2] Lu Tianbo, Guo Rui, Zhao Lingling, *et al.* A systematic review of network flow watermarking in anonymity systems [J]. International Journal of Security and Its Applications. 2016, 10(3): 129-138.

[3] Khan S, Gani A, Wahab A W A, *et al.* Network forensics: review, taxonomy, and open challenges [J]. Journal of Network and Computer Applications，2016, 66: 214-235.

[4] Pyun Y J, Park Y H, Wang Xinyuan, *et al.* Tracing traffic through intermediate hosts that repacketize flows [C]// Proc of the 26th IEEE International Conference on Computer Communications.Piscataway,NJ: IEEE Press, 2007: 634-642.

[5] Pyun Y J, Park Y H, Reeves D S, *et al.* Interval-based flow watermarking for tracing interactive traffic [J]. Computer Networks. 2012, 56(5): 1646-1665.

[6] Wang Xinyuan, Chen Shiping, Jajodia S. Network flow watermarking attack on low-latency anonymous communication systems [C]// Proc of the 28th IEEE Symposium on Security and Privacy. Washington DC: IEEE Computer Society, 2007: 116-130.

[7] Wang Xinyuan, Reeves D S. Robust correlation of encrypted attack traffic through stepping stones by flow watermarking [J]. IEEE Trans on Dependable and Secure Computing. 2010, 8(3): 434-449.

[8] Houmansadr A, Kiyavash N, Borisov N. Non-blind watermarking of network flows [J]. IEEE/ACM Trans on Networking. 2014, 22(4): 1232-1244.

[9] Houmansadr A, Kiyavash N, Borisov N. RAINBOW: a robust and

invisible non-blind watermark for network flows [C]// Proc of the 16th Annual Network & Distributed System Security Symposium. San Diego: Internet Society, 2009: 224-236.

[10] Houmansadr A, Borisov N. SWIRL: a scalable watermark to detect correlated network flows [C]// Proc of the 18th Annual Network & Distributed System Security Symposium. San Diego: Internet Society, 2011: 1-15.

[11] 张连成，王振兴，刘慧生. 网络流水印技术研究进展 [J]. 计算机科学，2011，38(11):7-11，42. (Zhang Liancheng, Wang Zhenxing, Liu Huisheng. Survey on network flow watermarking technologies [J]. Computer Science, 2011, 38(11): 7-11, 42.)

[12] Yu Wei, Fu Xinwen, Graham S, *et al.* DSSS-Based flow marking technique for invisible traceback [C]// Proc of the 28th IEEE Symposium on Security and Privacy. Washington DC: IEEE Computer Society, 2007: 7-21.

[13] Jia Weijia, Tso F P, Ling Zhen, *et al.* Blind detection of spread spectrum flow watermarks [J]. Security and Communication Networks. 2013, 6 (3): 257-274.

[14] Jia Weijia, Tso F P, Ling Zhen, *et al.* Blind detection of spread spectrum flow watermarks [C]// Proc of the 28th IEEE International Conference on Computer Communications. Washington DC:IEEE Computer Society, 2009: 2195-2203.

[15] 张连成，王禹，孔亚洲，等. 网络流水印安全威胁及对策综述 [J]. 计算机研究与发展, 2018, 55(8): 1785-1799. (Zhang Liancheng, Wang Yu, Kong Yazhou, *et al.* Survey on security threats and countermeasures of network flow watermarking [J]. Journal of Computer Research and Development, 2018, 55(8): 1785-1799.)

[16] Kiyavash N, Houmansadr A, Borisov N. Multi-flow attacks against network flow watermarks: analysis and countermeasures [J]. Columbia Law Review. 2012, 97: 1343-2470.

[17] Kiyavash N, Houmansadr A, Borisov N. Multi-flow attacks against network flow watermarking schemes [C]// Proc of the 17th USENIX Security. Berkeley: USENIX Association, 2008: 307-320.

[18] Zhang Liancheng, Wang Zhenxing, Wang Qinglong, *et al.* MSAC and multi-flow attacks resistant spread spectrum watermarks for network flows [C]// Proc of the 2nd IEEE International Conference on Information and Financial Engineering. Piscataway,NJ:IEEE Press, 2010: 438-441.

[19] Zhang Liancheng, Wang Zhenxing, Wang Yu, *et al.* Interval-based spread spectrum watermarks for tracing multiple network flows [C]// Proc of the 12th IEEE International Conference on Communication and Technology. Piscataway,NJ:IEEE Press, 2010: 394-397.

[20] 张连成，王振兴，孙建平. 基于时间间隔的扩频流水印技术 [J]. 计算机应用研究，2011，28(8): 3049-3053. (Zhang Liancheng, Wang Zhenxing, Sun Jianping. Interval-based spread spectrum watermarking scheme for tracing network flows [J]. Application Research of Computers, 2011, 28(8): 3049-3053.)

[21] 张璐，罗军舟，杨明，等. 基于时隙质心流水印的匿名通信追踪技术 [J]. 软件学报，2011，22(10): 2358-2371. (Zhang Lu, Luo Junzhou, Yang Ming, *et al.* Interval centroid based flow watermarking technique for anonymous communication traceback [J]. Journal of Software, 2011, 22(10): 2358-2371.)

[22] Paxson V. Growth trends in wide-area TCP connections [J]. IEEE Network，1994, 8(4): 8-17.

[23] Luo Xiapu, Zhang Junjie, Perdisci R, *et al.* On the secrecy of spread-spectrum flow watermarks [C]// Proc of the 15th European Symposium on Research in Computer Security. Athens: Springer, 2010: 232-248.